

Title: Functional consequence of positive selection revealed through rational mutagenesis of human myeloperoxidase.

Research Article

Noeleen B. Loughran^{1,2}, Sara Hinde³, Sally McCormick-Hill³, Kevin G. Leidal³, Sarah Bloomberg³, Sinéad T. Loughran⁴, Brendan O'Connor⁴, Ciarán Ó Fágáin⁴, William M. Nauseef^{3*}[†], and Mary J. O'Connell^{1,2*}[†]

¹Bioinformatics and Molecular Evolution Group, School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland

²Centre for Scientific Computing & Complex Systems modeling (SCI-SYM), Dublin City University, Glasnevin, Dublin 9, Ireland

³Iowa Inflammation Program, Department of Medicine, University of Iowa and Veterans Affairs Medical Center, Iowa City, Iowa 52241, USA

⁴School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland.

[†]Drs. Nauseef and O'Connell contributed equally to this work.

To whom correspondence should be addressed:

*William M. Nauseef, M.D.,
Iowa Inflammation Program and Department of Medicine,
Roy J. and Lucille A. Carver College of Medicine,
University of Iowa,
D160 MTF,
2501 Crosspark Road,
Coralville, IA 52231 (USA)

email: william-nauseef@uiowa.edu

Phone: +1 319 335 4278

Fax: +1 319 335 4194

*Mary J. O'Connell, PhD,
Bioinformatics and Molecular Evolution Group,
School of Biotechnology,
Dublin City University,
Glasnevin,
Dublin 9.
Ireland

email: mary.oconnell@dcu.ie

Phone: +353 1 700 5112

Fax: +353 1 700 5412

Running head: Positive selection and functional shift

Keywords: myeloperoxidase, animal peroxidase family, positive selection, protein evolution, Darwinian selection, functional shift.

Abbreviations: APF: 3'-(*p*-aminophenyl) Fluorescein, DB: Double Mutant, DMF: Dimethylformamide, EPO: Eosinophil peroxidase, ER: Endoplasmic Reticulum, HEK: Human Embryonic Kidney 293, HOCl: Hypochlorous Acid, H₂O₂: Hydrogen Peroxide, LPO: Lactoperoxidase, MHP: Mammalian Heme Peroxidase, ML: Maximum Likelihood, MPO: Myeloperoxidase, PDB: Protein data bank, PP: Posterior Probability, SEM: Standard Error of the Mean, TMB: 3,3',5,5'-Tetramethyl Benzidine, TPO: Thyroid peroxidase, TRI: Triple Mutant, WT: Wild Type.

The authors declare no conflict of interest.

Abstract

Myeloperoxidase (MPO) is a member of the mammalian heme peroxidase (MHP) multigene family. Whereas all MHPs oxidize specific halides to generate the corresponding hypohalous acid, MPO is unique in its capacity to oxidize chloride at physiologic pH to produce hypochlorous acid (HOCl), a potent microbicide that contributes to neutrophil-mediated host defense against infection. We have previously resolved the evolutionary relationships in this functionally diverse multigene family and predicted *in silico* that positive Darwinian selection played a major role in the observed functional diversities (2008, 8; DOI:10.1186/1471-2148-8-101). In this work we have replaced positively selected residues asparagine 496 (N496), tyrosine 500 (Y500) and leucine 504 (L504) with the amino acids present in the ancestral MHP and have examined the effects on the structure, biosynthesis, and activity of MPO. Analysis *in silico* predicted that N496F, Y500F or L504T would perturb hydrogen bonding in the heme pocket of MPO and thus disrupt the structural integrity of the enzyme. Biosynthesis of the mutants stably expressed in HEK cells yielded apoproMPO, the heme-free, enzymatically inactive precursor of MPO, that failed to undergo normal maturation or proteolytic processing. As a consequence of the maturational arrest at the apoproMPO stage of development, cells expressing MPO with mutations N496F, Y500F, L504T, individually or in combination, lacked normal peroxidase or chlorinating activity. Taken together, our data provide further support for the *in silico* predictions of positive selection and highlight the correlation between positive selection and functional divergence. Our data demonstrate that directly probing the functional importance of positive selection can provide important insights into understanding protein evolution.

Introduction

The relationship between positive selective pressure, as measured by nucleotide substitution processes and codon based models of evolution (Yang 1997, Yang et al. 2000), and functional divergence has rarely been investigated at both the genotypic and phenotypic level. Levasseur *et al.* studied the fungal lipase/feruloyl esterase A family to determine if a correlation existed between evolutionary, functional and environmental shifts (Levasseur et al. 2006). The signature of positive selection was detected across independent lineages of the multigene phylogeny. By integrating functional data from *in vitro* site-directed mutagenesis, they revealed that certain amino acids under positive selection were involved in the observed functional shift. The relationship between positive selection and protein functional shift has also been examined in antimicrobial peptides, with the conclusion that the conventional Dn/Ds ratio method to detect positive selection accurately identifies protein functional divergence (Tennessen 2008). Yokoyama *et al.* (2008) investigated the evolution of phenotypic adaptations using visual pigments in vertebrates, however, upon *in vitro* mutagenesis these sites revealed no significant influence on the adaptation of rhodopsin sensitivity, refuting the link between positive selection and functional divergence (Yokoyama et al. 2008). In contrast, Moury and Simon (2011) demonstrated how sites under positive Darwinian selection were involved in adaptive trade-offs between different fitness traits such as viral accumulation and transmissibility in *Potato virus Y* (Moury and Simon 2011). Studies involving the experimental validation or refutation of computational predictions such as detection of positive selection are challenging but provide us with far greater understanding of protein evolution. The need for experimental validation of these *in silico* predictions of positive selection has been clearly argued (Hughes 2008).

Previously, we identified a strong signature of positive selective pressure on the four main members of the mammalian heme peroxidase (MHP) family: myeloperoxidase (MPO), eosinophil peroxidase (EPO), thyroid peroxidase (TPO) and lactoperoxidase (LPO) (Loughran et al. 2008). We proposed that this positive selection contributed to the observed functional diversity in these enzymes (see Figure 1 for summary of this work). MPO (EC 1.11.1.7) is a homodimeric heme-containing protein found predominantly in the azurophilic granules of neutrophils, where its function is critical for optimal oxygen-dependent killing of ingested microbes (Johnson et al. 1987, Klebanoff 1970, 1991). All members of the MHPs participate in both peroxidation and halogenation cycles, whereby compound I produced by reaction with H_2O_2 catalyzes one- and two-electron oxidations, respectively. However, MPO is unique amongst the MHPs in its capacity to catalyze the two-electron oxidation of chloride at physiologic pH, thereby generating hypochlorous acid (HOCl). Thus, the unique chlorinating capacity of MPO is ideally tailored for its role in killing ingested microbes within neutrophils. We predicted that a small number of amino acid residues played an important role in the evolution of the unique capacity of MPO to chlorinate targets.

To test our hypothesis, we performed targeted mutagenesis followed by biochemical analyses to assess the functional consequences of replacing specific amino acids in MPO, an approach that we have employed successfully in the past to dissect the consequences of inherited mutations on the synthesis, structure, and function of MPO [Supplementary S1] (DeLeo et al. 1998, Goedken et al. 2007, Nauseef et al. 1996). The biosynthesis of mature MPO includes a series of critical structural modifications involving the incorporation of heme that leads

ultimately to the production of an active homodimer (see Figure 2 for a schematic of normal human MPO biosynthesis). In the endoplasmic reticulum (ER) the primary 80-kDa translation product, preproMPO, undergoes co-translational N-linked glycosylation to yield the enzymatically inactive apoproMPO (90-kDa). Subsequent insertion of heme into the peptide backbone of apoproMPO generates the active 90-kDa heme-containing precursor, proMPO, which exits the ER into the Golgi. Heme incorporation during MPO biosynthesis is a prerequisite for its activity and for proper proteolytic processing and trafficking to produce the mature homodimer in the azurophilic granule. The two identical monomers are linked by a disulphide bridge at position C319 resulting in mature dimeric MPO of approximately 150-kDa (Hansson et al. 2006), with each monomer consisting of a heavy and light subunit of 59-kDa and 13.5-kDa respectively.

Based on the resolved phylogenetic history of the MHPs (Figure 1) and positive selection identified in our *in silico* evolutionary study (Loughran et al. 2008), we mutated three sites (both *in silico* and *in vitro*), N496 (posterior probability (PP) = 0.999), Y500 (PP = 0.731) and L504 (PP = 0.970), to their more ancestral state, phenylalanine (F), F, and threonine (T), respectively. These positions were chosen based on: (i) the confidence score (PP) of being under positive selection from the *in silico* predictions (Loughran et al. 2008), (ii) their spatial relationship with the proximal heme ligand, His 502, and (iii) their proximity to R499 and G501, inherited mutations of which cause MPO deficiency (Goedken et al. 2007). Based on the resolved phylogenetic history of the MHPs (Figure 1), we mutated N496, Y500 and L504 sites *in vitro*.

We demonstrate that the substitutions N496F, Y500F and L504T, independently and in combination (*i.e.* double/triple mutants), disrupted normal MPO biosynthesis and severely compromised enzymatic activity. Our findings indicate that these residues are indeed closely associated with MPO-specific enzymatic function and support the hypothesis that the predicted residues positively selected in the MPO lineage contributed to its functional divergence from other MHP family members. Furthermore, these studies validate the experimental approach of directly testing *in vitro* the impact of *in silico* evolutionary predictions on protein function.

Materials & Methods

Reagents

Human embryonic kidney 293 (HEK) cells were obtained from American Type Culture Collection (Manassas, VA), ATCC CRL-1573. Mutagenesis kit was obtained from Stratagene, pcDNA3.1(-) Neo expression vector, G-418 sulphate antibiotic, the EnzChek® Myeloperoxidase (MPO) Activity Assay Kit and all tissue culture reagents were obtained from Invitrogen. [³⁵S]methionine (EasyTagEXPRESS Protein Labeling Mix, >37 TBq/mml) was obtained from Perkin-Elmer. Monospecific rabbit antibody against human myeloperoxidase was generated as noted previously (Nauseef et al. 1983). The SuperSignal® west pico chemiluminescent substrate was obtained from Thermo Fisher Scientific. All other reagents used in this study were obtained from Sigma-Aldrich.

Ancestral State Reconstruction, 3D Modeling and In Silico Mutational Analysis

In our previous study (Loughran *et al.* 2008), the phylogeny for the MHP multigene family of enzymes was fully resolved. Using this phylogeny and corresponding alignment, the ancestral node (amino acid level) of each clade on the phylogeny was reconstructed using the maximum likelihood (ML) marginal reconstruction algorithm implemented in PAML v3.15 (Yang 1997). Homology modeling and ancestral state *in silico* mutational analysis of the three MPO-specific positively selected sites (positions Asn496, Tyr500 and Leu504) were performed using SWISS-MODEL and DeepView v3.7 respectively (Arnold et al. 2006, Guex and Peitsch 1997). The effects on hydrogen bonding of the single mutations, N496F, Y500F and L504T, and of the double/triple mutations; DB1: N496F/Y500F, DB2: N496F/L504T, DB3: Y500F/L504T and N496F/Y500F/L504T, were assessed by *in silico* analysis as described previously (Loughran

et al. 2008). The human MPO structure (PDB accession code 1d2vC) was used as a homology model template. A summary table for N496F, Y500F and L504T is available as Supplementary S2).

In vitro Mutational Analysis

cDNA encoding normal human MPO (Johnson et al. 1987) was cloned into the expression vector pcDNA3.1(-). Neo and site directed mutagenesis was performed using the Stratagene QuickChange®II XL Site-Directed Mutagenesis Kit as per the manufacturer's protocol. Primers were designed to incorporate the three single (N496F, Y500F and L504T), the double mutants (N496F/Y500F, N496F/L504T, Y500F/L504T), and the triple mutant (N496F/Y500F/L504T). N496F forward primer: 5'-CGTCTTCACCT**TT**GCCTTCCGC-3'; reverse primer: 5'-GCGGAAGGCAA**AG**GTGAAGACG-3'. Y500F forward primer: 5'-CCTTCCGCT**TT**TGGCCACACCC-3'; reverse primer: 5'-GGTGTGGCC**AA**AGCGGAAGGC-3'. L504T forward primer: 5'-GGCCACACC**ACC**ATCCAACCC-3'; reverse primer: 5'-GGGTTGGATGG**T**GGTGTGGCC-3'. Letters in boldface indicate the specific nucleotide base changes required to alter the amino acid coded. Presence of the desired mutant and absence of unintentional mutations were verified by sequencing (Eurofins MWG Operon, London, U.K. and Integrated DNA Technologies, Iowa, U.S.A)

Stably Transfected Cell Lines

HEK cells were maintained in Dulbecco's modified Eagle's medium/Ham's nutrient mixture F-12 medium supplemented with 10 % (v/v) fetal bovine serum, 100 U/mL penicillin, 100 µg/mL streptomycin, 100 mM HEPES and 2 mM L-glutamine. HEK cells were then

transfected with mutant cDNA using the Qiagen® PolyFect Transfection Reagent and transfection procedure guidelines. Stable transfectants were selected using G-418 sulphate, as described previously (Goedken et al. 2007).

MPO Biosynthesis

Stable HEK transfections expressing wild type and mutant MPO were maintained in medium supplemented with 2 µg/mL hemin for 24 h prior to metabolic labeling and immunoprecipitation as described previously (Bulow et al. 2002, DeLeo et al. 1998, Goedken et al. 2007, Nauseef 1986, 1987, Nauseef et al. 1988, Nauseef et al. 1992, Nauseef et al. 1995, Nauseef et al. 1996, Nauseef et al. 1998). Briefly, cells were incubated in RPMI methionine-free medium supplemented with dialyzed fetal bovine serum, antibiotics and 2 µg/mL hemin for 1 h. Cells were pulse labeled with [³⁵S] methionine for 1 h and chased for 20 h by the addition of cold-methionine. Radiolabeled MPO-related protein in the cell lysate and medium was immunoprecipitated with rabbit polyclonal antiserum directed against human MPO. Samples were separated by SDS-PAGE followed by autoradiography and radioisotopically labelled MPO-related protein was quantified by densitometry using a PhosphorImager (Typhoon 9410, Amersham Biosciences).

MPO Activities

In addition to its classic peroxidase activity, MPO oxidizes chloride in the presence of H₂O₂ to generate the potent cytotoxic agent, HOCl. Consequently, we compared both peroxidase activity and chlorinating activity of mutant species and normal MPO. Nontransfected HEK cells and stably transfected lines expressing wild type and mutant MPO were maintained in medium

supplemented with 2 $\mu\text{g/mL}$ hemin for 48 h prior to cell recovery to assess MPO activity. To determine the relative amounts of MPO-related protein in lines expressing mutant MPO, cell lysates were separated by SDS-PAGE, blotted, and probed with monospecific anti-MPO antibody. MPO-related proteins in the immunoblot were quantified by densitometry using the Syngene Genegenius Bioimaging System and the amounts of MPO-related protein per cell were normalized relative to that in HEK cells stably expressing normal MPO, as done previously (Goedken et al. 2007).

(i) Peroxidase Activity Assay:

Cell pellets were solubilized in 0.01 % (v/v) Triton X-100/1X PBS at a density of 1×10^6 cells/18 μL and stored on ice. Peroxidase assay was performed in a waterbath held at 37 $^{\circ}\text{C}$. The reaction mixture contained 18 μL of cell lysate, 3.5 mL of assay buffer [1.4 mM tetramethyl benzidine (TMB), 8 % (v/v) dimethylformamide (DMF), 50 mM sodium acetate (pH 5.4)] and 2.1 μL of 0.49 M H_2O_2 (verified spectrophotometrically using $E_{240\text{nm}} = 43.6\text{M}^{-1}\text{cm}^{-1}$). The reaction was stopped after 3 min by the addition of 100 μL of 0.35 mg/mL catalase and 3.4 mL of ice-cold 0.2 M acetic acid. Absorbance at 655 nm was measured spectrophotometrically.

(ii) Chlorination Activity Assay – Technique 1 (performed for all non-N496F variants):

Cell pellets were resuspended in 1X PBS to a density of 1×10^6 cells/50 μL and sonicated on ice at 40 % amplitude for 30 sec with 6 sec pulses using a Branson Digital Sonifier (200W). Cleared lysate was collected by centrifugation for 10 min ($10,000 \times g$). The chlorination assay was performed in a microtiter plate using the EnzChek® Myeloperoxidase (MPO) Activity Assay Kit. Briefly, 50 μL of 2X 3'-(*p*-aminophenyl) fluorescein (APF) working solution was added to 50

μL of sample. The reaction mixture was then incubated in the dark at 37 °C for 5 min. Fluorescence intensity of each sample was recorded at 485 nm excitation and 530 nm emission on a Perkin Elmer luminescence spectrofluorometer.

(iii) Chlorination Activity Assay – Technique 2 (performed for all N496F variants):

MPO-specific chlorinating capacity was quantitated for N496F variants using a sensitive and specific assay for detection of HOCl-dependent generation of taurine monochloramine (Goedken et al. 2007). HEK cells, both wild-type (negative control) and transfectants were treated with diisopropylfluorophosphate (1 mM) for 20 minutes and sonicated for 10 seconds on ice. Cell sonicates were centrifuged (228 x g) to remove intact cells and the supernatant spun (10,800 x g) to recover the membrane-bound compartment containing normal or mutant MPO, as previously described (Goedken et al. 2007). The pellet was subjected to three sequential freeze-thaw cycles in phosphate buffered saline (without calcium or magnesium) with 0.3% cetyltrimethylammonium bromide and 1% protease inhibitor cocktail. The solubilized pellet was clarified by centrifugation (50,000 X g, 10 min, 4°C, in TL 120.2 rotor) and the clarified contents used in the assay as described (Dypbukt et al. 2005), with a standard curve for HOCl production generated using purified MPO (0-500 fmoles). Data are expressed as nmoles of HOCl produced per 10⁶ cell equivalents and all assays were performed in triplicate and each experiment was performed at least three times.

Results

Impact of mutation of positively selected sites on the structural integrity of MPO in silico

Performing detailed *in silico* site-directed mutagenesis, we previously assessed the impact of positively selected sites on the hydrogen bonding around the heme group in MPO (Loughran et al. 2008) and hence on the structural integrity of the enzyme [Supplementary S3-5]. Three positions, N496, Y500 and L504, are located up- and down-stream of the proximal heme ligand His 502. Based on our *in silico* mutagenesis of the positively selected residues in MPO, we predicted that the mutants N496F, Y500F and L504T were highly likely to compromise the heme-binding pocket of human MPO.

Position N496 shares a hydrogen bond with N587, which in turn is connected to the heme ligand, H502, via an additional hydrogen bond. Upon mutation of position 496 to phenylalanine (F), the 496-587 link is lost, potentially upsetting the structural integrity of the enzyme and indeed the 502-587 (heme ligand) bond (Supplementary S3). Data from the 3D structure of the enzyme indicate that Y500 covalently links R499 and G501, two residues that are critical for stability of H502, as inherited mutations R499C or G501S result in MPO deficiency (Goedken, McCormick et al. 2007). The proximal heme ligand, H502, is connected to R499 via a hydrogen bond and covalently links G501 and T503. Y500 shares putative hydrogen bonds with Y462, A497 and T503. By mutating position 500 to a phenylalanine, we predict that the hydrogen bond with Y462 would be lost (see Supplementary S3). Residue L504 is directly bound to T503 and shares hydrogen bonds with G501 and K556. Mutating leucine at position 504 to threonine (L504T) is likely to create an additional hydrogen bond with G501 (see Supplementary S3). All possible combinations of N496, Y500 and L504 mutated to their respective ancestral states

should not change the outcome of the above predictions; N496F would result in the loss of a hydrogen bond, Y500F would also result in the loss of a hydrogen bond and L504T in the formation of an additional hydrogen bond (Supplementary S3-5). Other non-covalent interactions and steric hindrances were not assessed, but may also result from these mutations and, in turn, further compromise the structural integrity of MPO. To test the validity of our predictions as to the impact of these mutations on the structure of MPO, we examined the biosynthesis and enzymatic activity of mutant proteins expressed in a heterologous system.

Effect of mutating positively selected sites on the biosynthesis of MPO

We created HEK cell lines stably transfected with normal, wild-type (WT) or mutant MPO [Single: N496F, Y500F, and, L504T. Double: N496F-Y500F (DB1), N496F-L504T (DB2), and, Y500F-L504T (DB3), and Triple: N496F-Y500F-L504T (TRI)]. All cell lines were biosynthetically radiolabeled with [³⁵S]-methionine for 1 h and chased for 0 and 20 h prior to immunoprecipitations of cell and culture medium with MPO antiserum.

HEK cells expressing normal MPO synthesized the 90-kDa precursors, apoproMPO and proMPO, which subsequently underwent proteolytic processing to yield mature MPO, represented by the appearance of the 59-kDa heavy subunit, as previously reported (Goedken, McCormick et al. 2007). Like transfectants expressing normal MPO, each of the mutant-expressing cell lines synthesized a 90-kDa precursor and 75-kDa intermediate species of MPO to varied degrees after pulse-labeling (Figure 3a-g). However, in contrast to the fate of normal MPO, mutant MPO precursors were not efficiently processed into mature enzyme. To assess the overall fate of MPO precursors in stable transfectants during the chase period, we directly

measured the amount of radioisotopically labeled MPO-related proteins and calculated the ratio of 90-kDa precursor to 59-kDa mature heavy subunit (90-kDa:59-kDa) at the end of the chase period as an approximation of the extent of processing (Table 1). For cells expressing normal MPO, the 90-kDa:59-kDa ratio was 0.79 ± 0.13 ($n = 10$). In contrast, the failure of mutant MPO precursors to be processed was best illustrated by the excess 90-kDa relative to mature MPO, with 90-kDa:59-kDa for five of seven mutants more than two-fold that of normally processed MPO (Table 1). Taken together, these data demonstrate that N496F, Y500F and L504T, individually or in combination, compromised efficient and normal processing of MPO precursors into mature subunits.

Effects of N496F, Y500F and L504T on peroxidase and chlorinating activity in MPO

Heme acquisition by apoproMPO to form proMPO is the rate limiting step in MPO biosynthesis, as chemical inhibition of heme synthesis or mutagenesis of critical residues arrests normal proteolytic processing (Castaneda et al. 1992, Nauseef et al. 1992, Pinnix et al. 1994). Consequently, we reasoned that the defective processing of MPO with mutations N496F, Y500F and L504T, individually and in combination, may reflect reduced heme acquisition, thus resulting in lower enzymatic activity of the mutant protein products. To assess the impact of N496F, Y500F and L504T on the enzymatic activities of MPO, we measured both peroxidase and chlorination activity of lysates from cells stably expressing mutant MPO. Enzymatic activities were normalized to the level of MPO-related proteins in each mutant, as judged by immunoblotting and subsequent densitometry, thereby allowing a comparison of relative specific activities of the various products.

There was a significant reduction in the peroxidase activity of each mutant relative to that of WT MPO, with variation in relative activity among the individual mutants (Table 2). Single mutations N496F and Y500F preserved peroxidase activity better than did the L504T mutant, whereas the loss of peroxidase activity among the double mutants and the triple mutant was very similar, with the exception of N496F/Y500F. Residual peroxidase activity indicates that each of the mutants incorporated heme, and it is thus reasonable to conclude that these mutations in the heme pocket did not completely inhibit the incorporation of heme. However, none of the mutants, single or in combination, had more than trace chlorinating activity relative to that of normal MPO, a depression far greater than the reduction in peroxidase activity (Table 2). Given the greater oxidation potential required for chlorination than for peroxidase activity, the discordance between the two enzymatic activities suggests that the mutations compromised the functional potential of the heme environment without fully inhibiting heme acquisition. Like members of the MHP family of proteins excepting MPO, the mutants were capable of supporting the single electron oxidation in peroxidase reactions, but not the two electron transfer necessary to chlorinate substrates.

Overall our results demonstrate that the residues under positive selection, namely N496, Y500, and L504, had important functional effects on the resultant protein. The analyses show that these residues were essential for: (i) the proper proteolytic processing of MPO precursors to mature MPO, (ii) stable acquisition of heme to support maximal peroxidase activity, and (iii) creation of a heme environment that supports the distinctive chlorination activity of normal MPO.

Discussion

Positive Darwinian selection is the process by which beneficial mutations in a population are retained and fixed, and it is considered synonymous with protein functional shift. In general, one of the resultant copies following gene duplication has increased freedom to explore mutational space while the other copy executes the original function of the gene (Lynch 2002, Ohta 1988a, b). The epistatic (compensatory) interactions between coevolving residues also influence the evolutionary trajectories of gene duplicates (Dean and Thornton 2007, Hayden et al. 2011). Mutations that prove beneficial are retained in this extra copy through the process of positive selection, and over time selective pressure can give rise to new functions through a process known as neofunctionalisation (Hughes 1999). Therefore, it follows that to a large extent, protein selection drives functional shift or protein diversification as well as protein specialisation within multigene families (Levasseur et al. 2006). Based on our recent *in silico* analysis of the biologically important MHP family of enzymes, we proposed that residues N496, Y500 and L504 were critical for the diversification of enzyme function in this family (Loughran et al. 2008) and were positively selected with respect to the unique capacity of MPO to oxidize chloride and thereby chlorinate substrates at physiologic pH.

To test the hypothesis that positive selection of these three residues was a driving force in the evolutionary diversification of MPO, we examined the impact of mutations N496F, Y500F, N496F-Y500F, L504T, N496F-L504T, N496F-Y500F-L504T and Y500F-L504T on biosynthesis and function of MPO. The pulse-chase and functional analyses of stably transfected cell lines expressing mutant forms of MPO revealed a profound effect on the cellular fate and activity of MPO. Although the biosynthesis of a 90-kDa MPO precursor proceeded normally in transfectants

expressing mutant MPO, subsequent proteolytic processing was impaired in all seven mutants. Failure to generate mature enzyme from precursor was most profound in the three double mutants, where the 90-kDa: 59-kDa ratio, a reflection of the efficiency of proteolytic processing, was ~ 4-fold greater in comparison with that seen in normal MPO. Given that formation of proMPO is a prerequisite for generation of mature MPO and that heme acquisition by apoproMPO results in proMPO, we reasoned that mutations at N496, Y500 and L504 compromised stable heme binding by mutant apoproMPO. In fact, cell lysates from transfectants expressing mutant MPO exhibited depressed peroxidase activity. It is noteworthy that L504T impaired peroxidase activity much more dramatically than did any of the other mutants generated, whereas there was relatively more 90-kDa MPO-related protein at 20 hours in L504T- than in Y500F-expressing cells. Taken together, these two observations suggest that the bulk of the 90-kDa MPO-related protein recovered from L504T cells was the enzymatically inactive precursor, apoproMPO.

It is also noteworthy that the mutants supported minimal to no chlorination, the enzymatic activity uniquely associated with normal MPO, despite the presence of substantial peroxidase activity in some of the mutants (*e.g.* N496F and Y500F). Residues N496, Y500, and L504 are in close proximity to H502, the heme-ligating residue in the proximal heme pocket that is conserved in all members of MHP protein family. Acquisition of heme by apoproMPO in the ER is required not only for MPO activity but also for proper proteolytic processing and targeting of the mature enzyme, as chemical inhibition of heme synthesis (Castaneda et al. 1992, Nauseef et al. 1992, Pinnix et al. 1994) or inherited defects in synthesis of normal proMPO (Nauseef et al. 1994, Nauseef et al. 1996) arrests MPO processing at the apoproMPO stage. Patients with

inherited mutations R499C or G501S, residues also in the proximal heme pocket, have MPO deficiency (Goedken et al. 2007) (see Supplementary S1). Although the heme group in all members of the MHP family is covalently bound to the protein backbone by ester bonds with conserved aspartate and glutamate residues, MPO is unique in having a third covalent bond between the 2-vinyl group of the heme and M409. This additional covalent linkage between heme and the protein backbone results in distortion of the heme ring, with secondary effects on the spectroscopic properties and, most importantly, oxidizing capacity of MPO. Thus, conformational features of the heme pocket in MPO uniquely enable it to support oxidation of chloride and the generation of the potent antimicrobial agent HOCl. The profound reduction in chlorinating capacity of all the mutants studied, even those with residual peroxidase activity, suggests that N496F, Y500F and L504T, like the inherited mutations at R499 and G501, altered the environment around H502 in MPO and compromised optimal activity by perturbing the wild type hydrogen bonding pattern in the heme binding pocket.

Taken together, these data have implications that apply both specifically to the MHP protein family as well as more generally to studies of protein functional shifts during evolution. Our studies support the concept that positively selected sites, N496, Y500 and L504 within the MPO protein, contribute to the observed protein functional shift in the MHP protein family. The impact on the peroxidation and chlorination activity caused by mutating these positions in MPO revealed the biological significance of the *in silico* predictions. The unique property of MPO to produce the potent oxidant HOCl was significantly disrupted following mutation of positively selected residues, and this loss of function suggests that these residues have accommodated the beneficial new function of chlorination activity and the production of HOCl in the MPO protein

lineage. Considered from a broader perspective, these studies demonstrate that biochemical and molecular approaches can be applied to test predictions generated *in silico*. Using targeted mutagenesis to probe predictions regarding evolutionary biology has a major role to play in the future elucidation of protein biology and evolutionary medicine.

Supplementary Material

Supplementary Tables:

Supplementary S1: Summary of mutations associated with MPO deficiency. Known causative and putative mutations identified by genotyping MPO deficient patients.

Supplementary S2: Summary of the effect of the mutations on hydrogen bonding within the myeloperoxidase structure. Mutation of positively selected residue in human MPO to more ancestral state (using human MPO model PDB accession code 1d2vC). Effect on H-bonds; “-” represents the loss of a putative hydrogen bond, whereas, “+” represents the gain of a putative hydrogen bond, respective to the order of their corresponding mutants.

Supplementary Figure Legends:

Supplementary S3: Effect of mutating individual positively selected sites on hydrogen bonding within the myeloperoxidase structure. (a) N496/F496 are shown in black and hydrogen bonding in grey. (b) Y500/F500 and (c) L504/T504 are shown in black, the heme ligand H502 in blue and hydrogen bonding in grey. Wild type is on the left and mutant on the right throughout. The “*” denotes a positively selected site.

Supplementary S4: Effect of double mutations on hydrogen bonding within the myeloperoxidase structure. (a) N496/F496 and Y500/F500 are seen in black, (b) N496/F496 and L504/T504 are seen in black, and (c) Y500/F500 and L504/T504 are seen in black. The heme ligand H502 can be seen in blue and hydrogen bonding in grey. Wild type is on the left and

mutant on the right throughout. The “*” denotes a positively selected site and the “^” denotes the unique covalent linkage site in MPO.

Supplementary S5: Effect of the triple mutation N496F/Y500F/L504T on hydrogen bonding within the myeloperoxidase structure. N496/F496, Y500/F500 and L504/T504 are seen in black, the heme ligand H502 in blue, and hydrogen bonding in grey. Wild type is on the left and mutant on the right. The “*” denotes a positively selected site and the “^” denotes the unique covalent linkage site in MPO.

Funding

This work was supported by the Irish Research Council for Science, Engineering and Technology (Embark Initiative Postgraduate Scholarship RS/2006/1016 to NBL). Dr Mary J. O’Connell is funded by Science Foundation Ireland, Research Frontiers Programme award EOB2673. Dr Brendan O’Connor is funded by Science Foundation Ireland’s strategic Cluster grant: Irish Separation Science Cluster. The Nauseef lab is supported by the National Institutes of Health grant AI 70958 (WMN) and with resources and use of facilities at the Iowa City Department of Veterans Affairs (VA) Medical Center, Iowa City, IA 52246. Travel was funded by the Benson Family and “The Orla Benson Award for Postgraduate Research”, School of Biotechnology, DCU. The Office for the Vice President for Research at Dublin City University is thanked for part-funding this research through the career start award (MJO’C).

Acknowledgements

The authors wish to acknowledge the SFI/HEA Irish Centre for High-End Computing (ICHEC) for the provision of computational facilities and support. We would like to thank Dr Patricia Johnson’s Viral Immunology Group, DCU for the use of their tissue culture facilities.

Literature Cited

- Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195-201.
- Bulow E, Nauseef WM, Goedken M, McCormick S, Calafat J, Gullberg U, Olsson I. 2002. Sorting for storage in myeloid cells of nonmyeloid proteins and chimeras with the propeptide of myeloperoxidase precursor. *J Leukoc Biol* 71: 279-288.
- Castaneda VL, Parmley RT, Pinnix IB, Raju SG, Guzman GS, Kinkade JM, Jr. 1992. Ultrastructural, immunochemical, and cytochemical study of myeloperoxidase in myeloid leukemia HL-60 cells following treatment with succinylacetone, an inhibitor of heme biosynthesis. *Exp Hematol* 20: 916-924.
- Dean AM, Thornton JW. 2007. Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet* 8: 675-688.
- DeLeo FR, Goedken M, McCormick SJ, Nauseef WM. 1998. A novel form of hereditary myeloperoxidase deficiency linked to endoplasmic reticulum/proteasome degradation. *J Clin Invest* 101: 2900-2909.
- Dypbukt JM, Bishop C, Brooks WM, Thong B, Eriksson H, Kettle AJ. 2005. A sensitive and selective assay for chloramine production by myeloperoxidase. *Free Radic Biol Med* 39: 1468-1477.
- Goedken M, McCormick S, Leidal KG, Suzuki K, Kameoka Y, Astern JM, Huang M, Cherkasov A, Nauseef WM. 2007. Impact of two novel mutations on the structure and function of human myeloperoxidase. *J Biol Chem* 282: 27994-28003.
- Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18: 2714-2723.
- Hansson M, Olsson I, Nauseef WM. 2006. Biosynthesis, processing, and sorting of human myeloperoxidase. *Arch Biochem Biophys* 445: 214-224.
- Hayden EJ, Ferrada E, Wagner A. 2011. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* 474: 92-95.
- Hughes AL. 1999. Adaptive evolution of genes and genomes. New York: Oxford University Press.
- . 2008. The origin of adaptive phenotypes. *Proc Natl Acad Sci U S A* 105: 13193-13194.
- Johnson KR, Nauseef WM, Care A, Wheelock MJ, Shane S, Hudson S, Koeffler HP, Selsted M, Miller C, Rovera G. 1987. Characterization of cDNA clones for human myeloperoxidase: predicted amino acid sequence and evidence for multiple mRNA species. *Nucleic Acids Res* 15: 2013-2028.
- Klebanoff SJ. 1970. Myeloperoxidase: contribution to the microbicidal activity of intact leukocytes. *Science* 169: 1095-1097.
- . 1991. Peroxidases in Chemistry and Biology. Boca Raton, Florida: CRC PRESS.
- Levasseur A, Gouret P, Lesage-Meessen L, Asther M, Record E, Pontarotti P. 2006. Tracking the connection between evolutionary and functional shifts using the fungal lipase/feruloyl esterase A family. *BMC Evol Biol* 6: 92.
- Loughran NB, O'Connor B, O'Fagain C, O'Connell MJ. 2008. The phylogeny of the mammalian heme peroxidases and the evolution of their diverse functions. *BMC Evol Biol* 8: 101.
- Lynch M. 2002. Genomics. Gene duplication and evolution. *Science* 297: 945-947.
- Marchetti C, Patriarca P, Solero GP, Baralle FE, Romano M. 2004a. Genetic studies on myeloperoxidase deficiency in Italy. *Jpn J Infect Dis* 57: S10-12.

- . 2004b. Genetic characterization of myeloperoxidase deficiency in Italy. *Hum Mutat* 23: 496-505.
- Moury B, Simon V. 2011. dN/dS-Based Methods Detect Positive Selection Linked to Trade-Offs between Different Fitness Traits in the Coat Protein of Potato virus Y. *Mol Biol Evol* 28: 2707-2717.
- Nauseef WM. 1986. Myeloperoxidase biosynthesis by a human promyelocytic leukemia cell line: insight into myeloperoxidase deficiency. *Blood* 67: 865-872.
- . 1987. Posttranslational processing of a human myeloid lysosomal protein, myeloperoxidase. *Blood* 70: 1143-1150.
- Nauseef WM, Root RK, Malech HL. 1983. Biochemical and immunologic analysis of hereditary myeloperoxidase deficiency. *J Clin Invest* 71: 1297-1307.
- Nauseef WM, Olsson I, Arnljots K. 1988. Biosynthesis and processing of myeloperoxidase--a marker for myeloid cell differentiation. *Eur J Haematol* 40: 97-110.
- Nauseef WM, McCormick S, Yi H. 1992. Roles of heme insertion and the mannose-6-phosphate receptor in processing of the human myeloid lysosomal enzyme, myeloperoxidase. *Blood* 80: 2622-2633.
- Nauseef WM, Brigham S, Cogley M. 1994. Hereditary myeloperoxidase deficiency due to a missense mutation of arginine 569 to tryptophan. *J Biol Chem* 269: 1212-1216.
- Nauseef WM, McCormick SJ, Clark RA. 1995. Calreticulin functions as a molecular chaperone in the biosynthesis of myeloperoxidase. *J Biol Chem* 270: 4741-4747.
- Nauseef WM, Cogley M, McCormick S. 1996. Effect of the R569W missense mutation on the biosynthesis of myeloperoxidase. *J Biol Chem* 271: 9546-9549.
- Nauseef WM, Cogley M, Bock S, Petrides PE. 1998. Pattern of inheritance in hereditary myeloperoxidase deficiency associated with the R569W missense mutation. *J Leukoc Biol* 63: 264-269.
- Ohta T. 1988a. Time for acquiring a new gene by duplication. *Proc Natl Acad Sci U S A* 85: 3509-3512.
- . 1988b. Evolution by gene duplication and compensatory advantageous mutations. *Genetics* 120: 841-847.
- Persad AS, Kameoka Y, Kanda S, Niho Y, Suzuki K. 2006. Arginine to cysteine mutation (R499C) found in a Japanese patient with complete myeloperoxidase deficiency. *Gene Expr* 13: 67-71.
- Pinnix IB, Guzman GS, Bonkovsky HL, Zaki SR, Kinkade JM, Jr. 1994. The post-translational processing of myeloperoxidase is regulated by the availability of heme. *Arch Biochem Biophys* 312: 447-458.
- Romano M, Dri P, Dadalt L, Patriarca P, Baralle FE. 1997. Biochemical and molecular characterization of hereditary myeloperoxidase deficiency. *Blood* 90: 4126-4134.
- Tennessen JA. 2008. Positive selection drives a correlation between non-synonymous/synonymous divergence and functional divergence. *Bioinformatics* 24: 1421-1425.
- Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13: 555-556.
- Yang Z, Nielsen R, Goldman N, Pedersen AM. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155: 431-449.

Yokoyama S, Tada T, Zhang H, Britt L. 2008. Elucidation of phenotypic adaptations: Molecular analyses of dim-light vision proteins in vertebrates. *Proc Natl Acad Sci U S A* 105: 13480-13485.

Table 1: Biosynthesis of normal and mutant MPO. Ratio of the densitometric calculations of immunoreactive MPO-related protein precipitated following pulse-chase analysis at 20 h (Mean \pm SEM).

Cell type	90kDa:59kDa
Normal	0.79 \pm 0.13 (n=10)
N496F	2.85 \pm 0.79 (n=3)
Y500F	2.06 \pm 0.36 (n=3)
L504T	1.65 \pm 0.39 (n=4)
DB1: N496F-Y500F	3.31 \pm 0.70 (n=3)
DB2: N496F-L504T	2.97 \pm 0.87 (n=3)
DB3: Y500F-L504T	3.22 \pm 0.56 (n=3)
TRI: N496F-Y500F-L504T	1.82 \pm 0.41 (n=3)

Table 2. Relative Specific Activity. Percentage peroxidase activity and chlorination of each mutant MPO with respect to wild type MPO, normalized to amount of MPO-related protein stably expressed by HEK cells (Mean \pm SEM, $n = 4$).

	Activity %	
	Peroxidation	Chlorination
MPO	100*	100**
N496F	55.2 \pm 1.2 ⁱ	0.00
Y500F	72.4 \pm 5.8 ⁱⁱ	0.00
L504T	1.1 \pm 1.1 ^{iv}	0.00
DB1: N496F-Y500F	46.4 \pm 1.7 ⁱⁱⁱ	0.23 \pm 0.03
DB2: N496F-L504T	23.6 \pm 1.7 ^v	0.33 \pm 0.03
DB3: Y500F-L504T	20.71 \pm 2.31 ^{vii}	0.00
TRI: N496F-Y500F-L504T	19.3 \pm 1.3 ^{vi}	1.00 \pm 0.06

* The value for normal MPO, represented as 100%, is 0.113 \pm .003 Δ A655.

** The value for normal MPO, represented as 100%, is 5.46 \pm 0.05 nmoles HOCl/million cell equivalents. The significance values are as follows:

ⁱ: $p < 0.0001$

ⁱⁱ: $p < 0.02$

ⁱⁱⁱ: $p < 0.0001$

^{iv}: $p < 0.0001$

^v: $p < 0.0001$

^{vi}: $p < 0.0001$

^{vii}: $p < 0.0001$

[Note: The significance values for mutants exhibiting trace levels of chlorinating activity were $p < 0.0001$]

Figure Legends

Figure 1: Phylogeny of the mammalian heme peroxidases.

Evolutionary relationship of the mammalian heme peroxidases (MHP); myeloperoxidase (MPO), eosinophil peroxidase (EPO), lactoperoxidase (LPO) and thyroid peroxidase (TPO), as resolved in Loughran *et al.* 2008. Bootstrap support values from 1000 replicates are shown. The star denotes the branches that were tested for signatures of positive selection. Number of sites under positive selection for each superfamily with their corresponding posterior probabilities (confidence scores) are shown to the right of each clade.

Figure 2: Normal myeloperoxidase biosynthesis.

The 80-kDa primary translation product (preproMPO) consists of a prepeptide, a propeptide region, and codons destined to become the small and large subunits of mature MPO. Cotranslational N-linked glycosylation of preproMPO yields the inactive 90-kDa precursor (apoproMPO). ApoproMPO associates with molecular chaperones in the ER, resulting in the incorporation of heme and thereby generating the active precursor (proMPO). Subsequently, proMPO exits to the Golgi for further processing and granule targeting. A short-lived 75-kDa intermediate lacking the propeptide region is subsequently cleaved into a two-subunit (13.5- and 59-kDa) monomer form. A 150-kDa homodimer (mature MPO) is formed with each monomer linked by a disulphide bond. Some of the immature MPO (90-kDa) undergoes modification of its oligosaccharide side-chains and is secreted constitutively.

Figure 3: Biosynthesis of wild type and mutant MPO. HEK cells stably expressing wild type (WT) or mutant MPO (single; N496F, Y500F, L504T, double; N496F-Y500F, N496F-L504T

and Y500F-L504T - denoted as DB1, DB2 and DB3 respectively, and the triple mutant; N496F-Y500F-L504T – denoted as TRI) were pulse-labeled with ^{35}S -cysteine/methionine and chased at 0 and 20 h intervals. Cell lysates at 0 and 20 h, and culture medium at 20 h (data not shown), were collected and MPO-related protein was immunoprecipitated. Immunoprecipitates were analysed by SDS-PAGE and autoradiography.